

研究データリポジトリ整備・運用ガイドライン

平成 31 年 3 月 29 日

国際的動向を踏まえたオープンサイエンスの
推進に関する検討会

目次

1. 研究データの保存・公開とデータリポジトリの整備・運用	1
2. 研究データリポジトリの定義及び役割	2
3. FAIR 原則	2
4. 信頼できる研究データリポジトリの整備・運用に関する要件	3
4.1 運営体制	3
4.1.1 研究データリポジトリのミッション	3
4.1.2 運営組織	3
4.1.3 運営計画	3
4.1.4 データポリシー	4
4.2 情報基盤	4
4.2.1 研究データリポジトリの ICT インフラ（狭義の情報基盤）	4
4.2.2 データ管理システム	4
4.2.3 データバックアップシステム	5
4.2.4 データ検索システム	5
4.3 人的基盤	6
4.3.1 研究データリポジトリにおける業務	6
4.3.2 運用スタッフに求められる技能・経験	6
4.3.3 運用スタッフの配置・育成	7

1 研究データの保存・公開とデータリポジトリの整備・運用

研究環境の ICT 化の急速な進展により研究データの電子化が促進され、研究データは爆発的に増大している。研究データを収集・蓄積して公開することにより、各学問領域の研究活動の迅速化や学際的な研究活動の進展、データ科学による新たな研究領域の創成が期待されている。一方、研究データの信頼性確保や研究不正防止の観点からも、研究データの一定期間の保存と適切な公開や共有・相互利用が求められている。これらを背景に、2013 年 6 月の G8 ロック・アーン・サミットにおいて「オープンデータ憲章」が合意されるなど、オープンサイエンスの推進に向けた国際的な議論や取組が行われている。こうした国際的なオープンサイエンスの推進の取組の中で、データ共有のための原則が検討され、国際的な認証機関が設立された。

このような国際的な潮流に対応すべく、我が国では 2014 年 12 月に内閣府において有識者の検討会が立ち上げられ、我が国のオープンサイエンスに対する基本姿勢・基本方針等を示した報告書「我が国におけるオープンサイエンス推進のあり方について」が 2015 年 3 月に取りまとめられた。その後、2016 年 1 月に閣議決定された「第 5 期科学技術基本計画」においても、我が国のオープンサイエンスに関する基本姿勢等が示されている。また、文部科学省では、学術情報委員会（第 8 期）において同様に検討が進められ、研究のエビデンスとなるデータを保存・公開することの意義とそのための具体的方策について示した審議まとめを 2016 年 2 月に行っている。

さらに、2018 年 6 月に閣議決定された「統合イノベーション戦略」において、今日の「知の源泉」の鍵はデータ・情報であるとの認識の下、オープンサイエンスのためのデータ基盤の整備を行う方針を打ち出している。具体的に講ずる施策として、国立研究開発法人（国研）におけるデータポリシー策定の促進等とともに、研究データを格納するリポジトリの整備等のため、文部科学省が「クラウド上で共同利用できる研究データの管理・公開・検索を促進するシステム」を開発すること、また、内閣府が国際認証基準等を参考に「リポジトリの整備・運用のガイドライン」を策定し、大学・国研等に同ガイドラインの適用を推奨することを掲げている。

本ガイドラインでは、データリポジトリの整備・運用を行う国研及び大学等が、国際的にも信頼され、研究データの保存・公開を適切に進めることができるよう、その整備・運用の指針を示す。データリポジトリの整備・運用にあたっては、各国研・大学等の機関における資金や人的リソースなども考慮しつつ進める必要がある。

なお、本ガイドラインは、リポジトリの国際認証基準の動向や上記の我が国における研究データの管理・公開・検索を促進するシステムの開発の進展等に応じて、必要な改訂を行う。

2 研究データリポジトリの定義及び役割

本ガイドラインでいうデータリポジトリとは、電子的データの保存・共有等を行うための広い意味の情報基盤であり、計算機基盤（狭義の情報基盤）のみならず、運営体制及び人的基盤を含む。本ガイドラインは、国研・大学等により整備されるものであって、とくに公的資金により得られた研究データを再利用できる形で保存・管理し、当該機関内または機関外へのデータ利用サービスに供するものを対象とする（以下「研究データリポジトリ」、または単に「リポジトリ」という。）¹。

本ガイドラインの対象とする研究データリポジトリには、データ駆動型研究開発を促進し、新たな「知の源泉」を創出するため、以下の役割が求められる。

- 当該分野内で標準的または相互利用に適したデータ形式の利用、データ再利用を促進する付帯情報としてのメタデータを整備すること等により、研究データの相互運用性²を確保し、研究データの共有（公開を含む）を図ること
- データ管理の水準を維持し、データへのアクセスを継続的に提供すること
- 国際的にも信頼が得られる形で、リポジトリの整備・運用を行うこと

3 FAIR 原則

データを共有するための基準となる国際的な原則として、「FAIR 原則（FAIR Data Principles）」³がある。「FAIR」は、「Findable（見つけられる）」「Accessible（アクセスできる）」「Interoperable（相互運用できる）」「Re-usable（再利用できる）」の略で、データ公開の適切な実施方法を表現しており、本原則に準拠したデータを作成する機運が国際的に高まっている⁴。

「FAIR 原則」の要件はデータ自体の相互運用性を意識して定められているが、その性質上、信頼性の高いデータリポジトリに求められる要件とも極めて親和性が高い⁵。

¹ リポジトリには、大学・研究機関の成果を保存・管理・公開することを主眼とする「機関リポジトリ」、分野における研究資源を保存して活動促進に資する「分野リポジトリ」、分野や機関を限定しない「汎用リポジトリ」などの分類がある。

² ここでいう相互運用性は「Interoperability」の訳語であるが、通常想定される計算機やソフトウェアを相互に接続して運用・運転できることとは異なり、データが相互に通用する、データ同士をあわせて用いる場合に有効に利用されることを指す。

³ データ共有の基準としての FAIR 原則：<https://doi.org/10.18908/a.2018041901>

⁴ 欧州委員会が、欧州の研究データ公開のための情報基盤である「European Open Science Cloud」について、「FAIR 原則」をデータ共有と管理の基礎に位置付けているなど。

⁵ <https://ipres2017.jp/wp-content/uploads/keynotel.pdf>

4 信頼できる研究データリポジトリの整備・運用に関する要件

研究データ利活用協議会の小委員会である「ジャパン・データリポジトリ・ネットワーク」(JDARN)では、「FAIR原則」及び国際的な認証基準(Core Trust Seal⁶による「信頼できるデータリポジトリを認定するための中核的な統一要件」)を踏まえ、信頼できる研究データリポジトリの整備・運用に関する要件の検討を行っている。本ガイドラインは、JDARNによる中間的な検討結果を参考に、以下のとおり、信頼できる研究データリポジトリの整備・運用に関する要件を示す。

4.1 運営体制

4.1.1 研究データリポジトリのミッション

研究データリポジトリには、研究データを適切に保存・管理し、提供する役割がある。研究データリポジトリには、保存・管理や継続的なデータアクセスの提供を可能とするようなミッションを定めて、公表することが求められる。研究データリポジトリの整備・運用に責任をもつ国研・大学等(以下「整備・運用機関」という。)は、機関内の研究データリポジトリを運営する組織(以下「運営組織」という。)に協力し、これを支援しなければならない。

4.1.2 運営組織

研究データリポジトリは、中長期的に(基本的に5年以上)安定して運用されることが求められる。このため、研究データリポジトリはその運営に関し、以下について配慮することが求められる。

- 永続的な組織によって運営されること
- 人的リソース、ITリソース、資金を十分に有すること
- 運用スタッフが継続的な研修と専門的な能力開発を受けられること

また、研究データの急速な変化に対応するため、研究データリポジトリの運営組織及び運用スタッフが専門家のガイダンスを受けられることが不可欠である。例えば、内部の専門家、外部諮問委員会、対象コミュニティから、研究データリポジトリの運用に対するフィードバックを受けられるようにする、国内外の相互運用性を低下させないため、定期的なデータ整備・管理に関する国際会議への参加を促す等のしくみが必要である。

4.1.3 運営計画

運営組織は、研究データリポジトリのミッションを達成するため、リポジトリの運営計画を策定し、その実施状況を把握し、必要に応じて改定を行わなければならない。整備・運用機関は運営組織を支援し、協力する必要がある。

⁶ <https://www.coretrustseal.org/>

また、データへの継続的なアクセスを保証する観点から、予期せぬ運営終了といった場合のデータの移転やデータ提供者へのデータの返却についても、その際の対応を予め示すことが望ましい。

4.1.4 データポリシー

研究データリポジトリは、その整備・運営にあたり、基本的なデータポリシー⁷について検討し、文書化しておくことが大切である。整備・運用機関がデータポリシーを策定している場合は、それに準拠しつつ、必要に応じて、研究データリポジトリ単位で、より具体的なデータポリシーを定めることも有用である。権利や免責事項に関しては、法律の専門家の助言を得ることを推奨する。

4.2 情報基盤

4.2.1 研究データリポジトリの ICT インフラ（狭義の情報基盤）

研究データリポジトリには、データ預け入れ先としての信頼性と持続性の確保が必要である。そのためには、研究データリポジトリは自らのサービスレベルを検討し、サービスレベルに応じた計算機・ネットワーク等の情報基盤（以下「ICT インフラ」という。）を確保・運用する必要がある。研究データリポジトリの ICT インフラに関して、検討すべき項目を以下に挙げる。

- 安定的な財源及び安定的な運用形態の確保
- 研究データリポジトリのサービスレベルの定義
- 研究データリポジトリのサービスレベルに合わせた機密性、可用性、完全性の確保
- データ保存領域の拡張性及び移行可能性の確保

4.2.2 データ管理システム

研究データリポジトリは、提供されたデータの受入・保管、メタデータの管理・保管及びデータの変更を含むデータキュレーションを実施するとともに、その管理等を行うためのデータ管理システムを確保し、運営する必要がある。

データ管理システムは、データの完全性を確保するための機能を備え、リポジトリのデー

⁷ データポリシーに含まれる項目の例を以下に挙げる。

- リポジトリが受け入れるデータの範囲（分野、種類、品質、ファイルフォーマット、サイズ）、メタデータ（形式、スキーマ）、データ提供者の要件
- データの公開・非公開及び猶予期間（エンバゴ）
- データの帰属・知的財産の取り扱い
- 倫理規定
- データの品質管理（評価方法、責任）
- データの利用条件（ライセンス）
- データの保存（保存期間、バックアップ方針、不変性検証）
- データの削除（データ削除の条件、判断基準）

その他、「国立研究開発法人におけるデータポリシー策定のためのガイドライン」も参照。

タポリシーに合致したものでなければならない。

さらに、データ管理システムは、「FAIR 原則」を実現するために以下の機能を備えていることが望ましい。

- データ及びメタデータに関連付いた永続識別子⁸の管理（「見つけられる (Findable)」及び「アクセスできる (Accessible)」)
- 可能な限りオープン⁹なフォーマットによるデータの管理（「相互運用できる (Interoperable)」及び「再利用できる (Re-usable)」)
- データ及びメタデータのバージョンの管理（「再利用できる (Re-usable)」)
- 分野標準のリッチなメタデータ形式、あるいは分野内でできるだけ広く用いられる形式によるメタデータの管理（「相互運用できる (Interoperable)」及び「再利用できる (Re-usable)」)
- 汎用的なメタデータ形式によるメタデータの管理（「見つけられる (Findable)」)
- データのライセンス情報の管理（「再利用できる (Re-usable)」)

4.2.3 データバックアップシステム

研究データリポジトリは、保管するデータの破損や紛失に備えて、研究データのバックアップの取得・保管、及びデータのリストアを実施するためのデータバックアップシステムの確保を行う必要がある（外部機関の活用を含む。）。ここでいうシステムは、計算機システムによるストレージのバックアップのほか、制度・組織間スキーム、その他幅広くデータ保存に有益なしくみ全体を含む。

バックアップの方法や頻度は、リポジトリのデータポリシーに従う必要がある。また、災害等による障害を考慮し、遠隔地への複製機能を有することが望ましい。データバックアップシステムの構築に当たり検討すべき項目を以下に挙げる。

- バックアップ実行頻度
- 保持方式（フル/差分/増分）
- 保持リビジョン数
- バックアップ及びリカバリ手順
- 遠隔バックアップの有無

4.2.4 データ検索システム

研究データリポジトリは、保管するデータまたはそのメタデータに対してアクセス可能なデータ検索システムを確保し、運用しなければならない。

⁸ 永続識別子の例：

- Digital Object Identifier (DOI)
- ORCID
- Global Research Identifier
- Persistent Uniform Resource Locator (PURL)
- Identifiers.org

⁹ <https://opendefinition.org/od/2.1/ja/>

データ検索システムは、「FAIR原則」を実現するために以下の機能を備えていることが望ましい。

- データ及びメタデータに関連付いた永続識別子の表示
- 機械判読可能なメタデータ及び永続識別子の提供
- データアクセス手法の表示
- データのライセンス及び引用方法の表示
- データ及びメタデータの版または変更履歴の表示

4.3 人的基盤

4.3.1 研究データリポジトリにおける業務

研究データリポジトリには安定した運営組織とともにそれを担う人材が必要である。運用スタッフは、以下の3つの業務を円滑に行うことが求められる。

- 管理的業務：リポジトリのミッションに則した企画・運営、ポリシーの作成、広報・コミュニケーション活動など
- コンテンツ運用業務：データ・メタデータの受入・管理、データ提供やキュレーション（受入データの前処理、クリーニング、整形、管理）など
- 計算機システム運用業務：サーバーやストレージなど、計算機基盤の構築や運用・保守や障害復旧、セキュリティ対策、設備の管理など

現状のリポジトリの運営においては、組織の規模や形態により、運用スタッフの配置のあり方が異なっている。複数の業務を一人が兼務している場合、関連業務を複数の機関・部門・部署で分担している場合、業務の一部を外注している場合がある。

リポジトリを設置している整備・運用機関において、リポジトリ運営業務の位置付けが明確でなく、その重要性がその機関内や関連機関内で十分認識されていないことが、リポジトリ運営の不安定要因の一つとなっている。

整備・運用機関は、リポジトリの運用業務に多様な専門性を求められることを認識し、運用スタッフの配置・専門性の育成、それらにかかる財源等を措置することが求められる。

4.3.2 運用スタッフに求められる技能・経験

研究データリポジトリの業務に関わる人材の技能は、業務ごとに多岐にわたる。

前節で挙げた3つの業務に関わる人材に必要と考えられる主な技能として、以下が挙げられる。

- 管理的業務：リポジトリ運営全般（経営、財務、法律、ポリシー、各種契約、データ受入審査、データライセンス、コンプライアンス等）の知識・経験、リポジトリに関するコンサルタントや広報の技能・経験
- コンテンツ運用業務：データキュレーション、データ品質管理、メタデータ管理、永続識別子の付与、データ公開やデータカタログへの登録、個人情報管理、データ・メタデータに関するコンサルタントの技能・経験に加えデータ提供者や利用者とのコミュニ

ケーション力・経験⁹

- 計算機システム運用業務：ハードウェア・ソフトウェア・ネットワーク管理や障害対応、セキュリティ対策に関する技能・経験

また、研究データが有効に利活用されるために、リポジトリの適切な運営に加え、データ分析支援を行うデータサイエンティスト等が配置されることが望ましい。

4.3.3 運用スタッフの配置・育成

研究データリポジトリの運用スタッフの配置にあたり、一人のスタッフが兼務する場合もあるが、他部門のスタッフの活用も含め、複数のスタッフで分担出来る体制を考慮すべきである。そのためには、運用スタッフに加え、研究データリポジトリの業務に関連する他部門のスタッフ（技術職員、図書館職員、リサーチアドミニストレータなど）に対し、必要な情報と研修の機会を提供する必要がある。

スタッフが幅広い関連知識を習得できるよう、研究データ管理（Research Data Management）の基本知識を学べるオンライン教材の活用や受講、また関連する国際会議参加を推奨するなど、スタッフに対して適切な研修の機会を提供する必要がある。なお、研究データ管理に関する教材¹⁰が国内でも開発・公開されている。

また、スタッフが必要とする知識や情報を共有できるコミュニティに参加したり、新たにコミュニティを構築することで、スタッフのスキルアップに加え、リポジトリの相互運用性やデータの利活用に必要な情報共有等を行うことができる。

⁹ データ専門職（データライブラリアン、データキュレーター）の現状：

2016年に公表された科学技術・学術審議会 学術分科会 学術情報委員会による審議のまとめ (http://www.mext.go.jp/b_menu/shingi/gijyutu/gijyutu4/036/houkoku/1368803.htm)では、大学等による人材の育成及び確保のための取り組みとして、

- 技術職員、URA及び図書館職員を中心としたデータ管理体制を構築する
- データキュレーター等を育成するシステムを検討し推進する
- 戦略的な人的資源配分を検討し、新たな専門人材の確保に努める

の3点が挙げられている。データの専門職としてデータライブラリアン、データキュレーター等の名称が現れてきており、各大学、研究機関でデータ管理、データ公開を担ってきた人材（技術職員、リサーチアシスタント、大学図書館員等）を認知し、同名称のもと再定義する動きもある。

¹⁰ 2017年度にオープンアクセスリポジトリ推進協会（JPCOAR）及び国立情報学研究所（NII）によって開発・公開された、以下のオンライン教材（MOOC）がある。

オープンサイエンス時代の研究データ管理：<https://www.nii.ac.jp/service/jmooc/rdm/>

国際的動向を踏まえたオープンサイエンスの推進に関する検討会の開催について

平成 29 年 11 月 22 日（制定）

平成 31 年 2 月 26 日（改定）

内閣府 政策統括官（科学技術・イノベーション担当）決定

1. 趣旨

国際動向を踏まえたオープンサイエンス推進のための方策等について検討し、本政策分野における国際プレゼンスの向上、及び国内施策の充実を図るため、「国際的動向を踏まえたオープンサイエンスの推進に関する検討会」（以下、「検討会」という。）を開催する。

2. 構成

- （1）検討会は、総合科学技術・イノベーション会議議員及び外部有識者により構成し、政策統括官（科学技術・イノベーション担当）が開催する。
- （2）政策統括官（科学技術・イノベーション担当）は、構成員の中から座長を依頼する。
- （3）検討会は、必要に応じ、関係者の出席を求めることができる。

3. ワーキンググループ

検討会は必要に応じて、本検討会の下にワーキンググループを設置することができるものとする。

4. 公開

検討会は原則として公開する。

5. 検討会の庶務

検討会の庶務は、関係府省等の協力を得て、政策統括官（科学技術・イノベーション担当）付で処理する。

国際的動向を踏まえたオープンサイエンスの推進に関する検討会 構成員
(3月29日現在)

【外部有識者】

(座長)

引原 隆士 国立大学法人 京都大学 図書館機構長・附属図書館長

(副座長)

喜連川 優 大学共同利用機関法人
情報・システム研究機構 国立情報学研究所 所長

(構成員)

有川 節夫 放送大学学園 理事長、国立大学法人 九州大学 前総長

家 泰弘 独立行政法人 日本学術振興会 理事

小賀坂 康志 国立研究開発法人 科学技術振興機構 知識基盤情報部長

川村 光 一般社団法人 日本物理学会 会長

黒川 顕 大学共同利用機関法人 情報・システム研究機構
国立遺伝学研究所 生命情報研究センター 教授

末吉 亙 潮見坂総合法律事務所 弁護士

関口 智嗣 国立研究開発法人 産業技術総合研究所 理事 情報・人間工学領域長

高木 利久 国立大学法人 東京大学 大学院理学系研究科 教授

谷藤 幹子 国立研究開発法人 物質・材料研究機構 統合型材料開発・情報基盤部門
材料データプラットフォームセンター センター長

林 和弘 文部科学省科学技術・学術政策研究所
科学技術予測センター 上席研究官

村山 泰啓 国立研究開発法人 情報通信研究機構 戦略的プログラムオフィス研究統括

【総合科学技術・イノベーション会議】

上山 隆大 有識者議員

山極 壽一 有識者議員

-第2章 知の源泉- （2）オープンサイエンスのためのデータ基盤の整備

○目指すべき将来像

- ・国益や研究分野の特性等を踏まえて、オープン・アンド・クローズ戦略¹を考慮し、サイバー空間上での研究データ²の保存・管理に取り組み、諸外国の研究データ基盤とも連携して巨大な「知の源泉」を構築し、あらゆる者が研究成果を幅広く活用
- ・その結果、所属機関、専門分野、国境を越えた新たな協働による知の創出が加速

○目標

＜リポジトリ³の整備及び展開＞

- ・機関リポジトリ⁴を活用した研究データの管理・公開・検索を促進するシステムを開発し、2020年度に運用開始

＜研究データの管理・利活用についての方針・計画の策定等＞

- ・研究成果としての研究データの管理・利活用のための方針・計画の策定⁵を促進
- ・これらの方針・計画に基づき公的資金による研究データについて、機関リポジトリを始めとするデータインフラで公開を促進
- ・公的資金による研究成果としての研究データについては、データインフラを通して機械判読可能性と相互運用性を確保するとともに、公開する研究データについては諸外国の研究データ基盤との連携を促進

＜人材の育成及び研究データ利活用の実態把握＞

- ・研究データの利活用を図るため、研修教材の活用を促進するとともに、実態把握を行いながら、研究者や研究支援職員の意識を向上

○目標達成に向けた主な課題及び今後の方向性

- ・機関リポジトリにおける研究論文以外の研究データの登載や、研究データの管理・利活用の方針・計画の策定が進んでいないなど、取組が不十分であり、研究者のデータ管理・利活用の意識や基本的な考え方についての認識も低い
- ・内閣府（科技）は、国際認証基準等に基づくリポジトリの整備・運用のガイドライン及び国研におけるデータポリシーの策定を促進するためのガイドラインを2018年度に策定
- ・研究データの特性等を踏まえて研究データを保存・公開するためのリポジトリの整備や研究データの管理・利活用のための方針・計画の策定を促進し、データインフラを通じた機械判読可能性と相互運用性の確保、諸外国の研究データ基盤との連携を促進
- ・研究者や大学・国研等における現状・取組等についての調査・分析を行い、研究者等の意識向上等に資する方策を検討

¹ データの特性から公開すべきもの（オープン）と保護するもの（クローズ）を分別して公開する戦略。

² 研究成果（論文等）の根拠となるものを含む。

³ データインフラのうち、電子的な知的生産物の保存や発信を行うためのインターネット上のアーカイブシステム。

⁴ リポジトリのうち、大学・国研等が管理するもの。我が国では主に大学が管理している。

⁵ 研究データの管理・利活用のための方針については国研が2020年度末までに策定、計画については競争的研究費による研究実施者が策定することを要請する制度を2021年度予算における公募までに、各府省・研究資金配分機関において導入。

① イノベーションにおけるオープンサイエンスのためのデータ基盤の必要性・重要性

活版印刷技術の登場により、本や学術ジャーナルの大量印刷が可能となって以降、物流や知的財産制度の整備とともに、よりオープンな知的基盤が構築された。その結果、多くのイノベーションが創出され、人類の発展をもたらした。

これまで大量印刷と物流による知的基盤が支えてきた科学は、ICTの発展により、サイバー空間が支える科学へと大きく変容し、学界、産業界、市民等あらゆる者がサイバー空間にある研究データを利活用し、協働によって知の創出をするというオープンサイエンスが進展してきている。このような社会の変化に応じて、新たな制度を整備しつつ、研究データの取扱いについての対応方針や運用を再定義することが求められる。

したがって、今後も我が国の研究や産業をますます発展させるべくイノベーションを創出するためには、社会インフラとして、オープンサイエンスのためのデータ基盤の構築が必要である。

② 現状認識

現在、ICTを活用した研究データの利活用による科学研究の変容と新しいイノベーション基盤づくりの議論が国際的に進展している。

例えば、米国では、研究資金配分を行う 22 の連邦政府機関全てがそれぞれの所管分野や所管する政府研究機関における研究成果の利活用のための計画を作成済みである。また、EUでは、既存の研究データの流通基盤の統合を進めている。さらに、国際コンソーシアムにおいて、研究者・技術者等による研究データ流通に係る国際標準等についての議論も進んでいる。

我が国では、機関リポジトリは世界でも最多⁶であるが、研究論文以外の研究データの登載が進んでいないこともあり、機関リポジトリを活用した研究データを管理・公開・検索するプラットフォームを構築するためのシステム開発⁷が進められている。

また、研究データを管理・利活用する上で欠かせないデータポリシー⁸を策定した国研が2017年までに2法人⁹にとどまるなど、研究分野別・組織別の特性を考慮したデータポリシーの策定が遅れており、研究データの管理・利活用が進んでいない。

さらに、データマネジメントプラン¹⁰の策定の要請など、競争的研究費による研究において研究実施者に適切な研究データの管理・利活用を促す仕組みの導入が十分ではなく、研究者のデータ管理・利活用の意識や基本的な考え方についての認識も低い状況にある。

このままオープン・アンド・クローズ戦略を検討せずに研究データの公開が進み、我が国の産業振興等のために優先的に研究データを利活用できる機会を失い、他国の企業等が先んじて商業化等に利活用することがないよう、データポリシーやデータマネジメントプランの策定を急ぐ必要がある。

⁶ 各国の機関リポジトリ数は日本 744、米国 500、イギリス 255、ドイツ 244 と続く（N I I 調べ）。

⁷ 開発主体はN I I。

⁸ 研究データの管理・利活用についての方針を組織として取りまとめたもの。

⁹ J A M S T E C、N I E S の 2 法人。

¹⁰ 競争的研究費による研究実施者が策定する研究データの管理・利活用についての計画。

③ 今後の方向性及び具体的に講ずる主要施策

研究データを我が国のデータインフラから公開できるよう、主として機関リポジトリを対象としたシステム開発や、国際認証基準¹¹等を参考にしたリポジトリの整備・運用を進める。

その上で、研究分野の特性等を踏まえたオープン・アンド・クローズ戦略を考慮したデータポリシーやデータマネジメントプランの策定を促進し、これらに基づく研究データの管理・公開等を促進するとともに、公的資金による研究成果としての研究データについては、データインフラを通して機械判読可能化を促進する。

さらに、これらの取組が大学・国研等で適切に行われるよう、研究データの管理や公開・共有に従事する研究者等の意識向上や基礎的な知識の習得のための取組や、研究者や大学・国研等における現状、取組等についての調査・分析を行い、研究者等の意識向上等に資する方策を検討する。

i) リポジトリの整備及び展開

【科技、文】

- ・文部科学省が主体となり、機関リポジトリを活用しクラウド上で共同利用できる研究データの管理・公開・検索を促進するシステムを開発し、2020年度に運用開始
- ・文部科学省が主体となり、全文データベース¹²に登載された論文と識別子¹³を付与した研究データを紐づけ、管理・公開するシステムの開発を2018年度中に検討
- ・国際認証基準等に基づくリポジトリの整備・運用のガイドライン（公開データの検索可能化、諸外国の研究データ基盤との相互運用性等を含む。）を内閣府（科技）が策定し、大学・国研等にガイドラインの適用を推奨
- ・ドメイン語彙策定への関係機関の参画等を通じて分野間データ連携基盤と連携

ii) 研究データの管理・利活用についての方針・計画の策定等

【内閣官房、科技、食品、総、文、厚、農、経、国、環、防】

- ・内閣府（科技）は、国研におけるデータポリシーの策定を促進するためのガイドラインを2018年6月までに策定
- ・国研は、研究分野の特性、国際的環境、産業育成等に配慮し、必要に応じてオープン・アンド・クローズ戦略を取り入れ、データポリシーを策定¹⁴
- ・競争的研究費制度の目的、対象等を踏まえ、大学・国研・企業等の研究実施者がデータマネジメントプラン等のデータ管理を適切に行う仕組み¹⁵を、各府省・研究資金配分機関が所管の競争的研究費制度に導入¹⁶（ガイドライン策定や公募要領改訂等）
- ・データポリシー・データマネジメントプランに基づく公的資金による研究データの管理・公開等を促進し、公的資金による研究成果としての研究データについては、データインフラを通して機械判読可能化を促進

¹¹ ここでは、2017年9月に設立されたリポジトリの国際的な認証機関であるCTS（Core Trust Seal）が定めた認証基準を指す。

¹² 論文等の書誌（タイトル、著者名、発行年等）のみでなく、全文を登載し、閲覧のため提供する情報の集まり。

¹³ 研究成果に対し、それを一義的に識別し、国際的に通用するデジタル情報。

¹⁴ 策定法人は、2017年末時点で2法人であるが、2020年度末までに24法人（国研のうち、研究資金配分機関であるAMED、JST、NEDOを除く。）全てでの策定を目指す（法人内の特定の研究センター・部門等の範囲で策定した場合を含む。）。

¹⁵ データマネジメントプランの策定、指定したリポジトリでの公開、データ作成者名等の報告書等での明示等。

¹⁶ 導入府省・研究資金配分機関は、2018年度当初時点で4であるが、2021年度予算における公募までに14府省・機関全てでの導入を目指す（制度内で特定のプログラム・事業等で導入した場合を含む。）。

iii) 人材の育成及び研究データ利活用の実態把握

【科技、文】

- ・2017年度に開発・公開された基礎的な能力開発プログラム¹⁷について、専門性を高めた拡張版の開発・公開及び関係機関における受講の促進
- ・大学・国研等のデータポリシー等作成状況、リポジトリ整備状況、研究データ利活用の優良事例、研究者の公的資金による研究データの公開状況や利活用上の障壁・促進要因等の調査を2019年度から実施¹⁸

¹⁷ 主に大学の研究支援職員が研究データの管理に関する基礎的な知識の習得を目的とし、オープンアクセスリポジトリ推進協会が開発したもの。

¹⁸ N I S T E P「研究データ公開と論文のオープンアクセスに関する実態調査[調査資料-268]」(2017年12月)では、科学技術の専門家(回答者は約1400名)を対象として、先行的な調査を実施している。